

Moving from Katana to Gadi High Performance Computing

John Zaitseff

October 2024

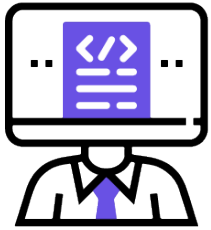
ResTech Compute | Community | Data



UNSW
SYDNEY

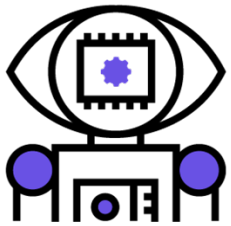
Research Technology Services (ResTech)

Compute



- High Performance Computing: Katana, Gadi, ...
- Cloud computing: Amazon AWS, Microsoft Azure, Google
- Some code and algorithm support

Data



- Storage: Data archive, Intersect, GitHub, encryption, ...
- Tools: REDCap, Qualtrics, Globus, ERICA, ...
- Data management support: grants, ethics applications



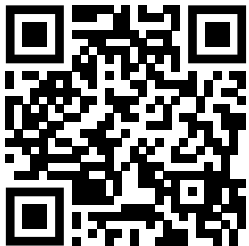
Research Technology Services (ResTech)

Community



- Over 50 free training courses: Linux, Python, R, AI/ML, NVivo, ...
- [Drop-In Forum](#): informal, peer-to-peer and experts
- Events and networking: conferences, seminars, videos, ...

Contact Us!



- Website: <https://unsw.sharepoint.com/sites/Restech>
- Email: restech@unsw.edu.au
- Location: [Level 3, June Griffith Building \(F10\)](#)

HPC resources for UNSW researchers

Katana

- Small HPC cluster at UNSW (165 nodes, 7384 CPU cores, 73 GPUs)
- Ideal for beginner to intermediate users (interactive use encouraged)
- Buy-in scheme and/or use nodes purchased by ResTech

Gadi

- National facility from NCI (4997 nodes, 260,760 CPU cores, 692 GPUs)
- Ideal for intermediate to advanced users (interactive use discouraged)
- National Merit Allocation Scheme or UNSW scheme (~200M SU per year)

HPC resources compared

	Katana	Gadi
Nodes; CPU cores; GPUs	165; 7384; 73	4997; 260,760; 692
Login nodes; Data mover nodes	3; 1	10; 6
Storage capacity	~ 12 PiB	~ 90 PiB (450 GiB/s and 960 GiB/s)
Number of users	~ 2500	~ 9000
Queue lengths (today)	274 running, 1271 queued	2164 running, 1512 queued
Resource limits	Depends on buy-in	Depends on allocation
Cost of each job	No additional cost	Depends on queue and resources
Home directory quota	15 GiB	10 GiB
Shared scratch directory quota	128 GiB by default	1 TiB by default
On-node fast scratch (JobFS)	No JobFS, but shared local scratch	100 MiB by default, up to 400 GiB
Global data directory	None	As requested by project
Number of applications; versions	596; 780	256; 967

When should I move?

From Katana to Gadi

- Your jobs have become too big for Katana: multi-node / MPI, lots of jobs
- Your jobs are *not* interactive and are production-ready
- Need to collaborate with non-UNSW researchers
- Need to access public / national data sets (compute where your data lives)

From Gadi to Katana

- Allows workflow experimentation without wasting NCI allocations
- Dedicated hardware resources purchased by your group
- Access to bespoke architectures (e.g., ARM-based Grace Hopper GPU node)

Getting started

1. Create an account at NCI (~10 minutes)

- <https://my.nci.org.au/mancini/signup/>
- Use your official UNSW email address
- Join an existing project if you know the project code (e.g., **aa00**)

2. Create a project (~1 hour)

- <https://my.nci.org.au/mancini/project/propose/>
- Must be done by an academic staff member (Lead Chief Investigator)
- Can be done later as needed

Getting started, steps 3 and 4

3. Submit resource requests for your projects

- National Merit Allocation Scheme (NCMAS): August–September each year
- UNSW Scheme: November–December each year
- Full details at <https://unsw.sharepoint.com/sites/Restech/SitePages/NCI.aspx>

4. Copy job scripts and data

- On Katana, use **rsync** with a destination of `gadi-dm.nci.org.au`
 - e.g., `rsync -vaSH ./job1/ jjz561@gadi-dm.nci.org.au:/home/561/jjz561/job1/`
- Can also use Globus: <https://app.globus.org/>

Getting started, steps 5 and 6

5. Modify job scripts as needed

- Change directory references (home, scratch, JobFS/local scratch)
- Change **module load / unload** invocations
- Change **#PBS** directives

6. Submit your jobs

- Determine resources needed: use more/less processor cores? memory? time?
- Determine the correct queue to use: [queue limits](#), [characteristics](#)
- Submit using **qsub -q *queue_name***

Directories and user/project names

	Katana	Gadi
User name	Your zID (e.g., z1234567)	Letters/digits (e.g., zz1234)
Project names	Letters/digits (e.g., group123)	Letters/digits (e.g., aa00)
Home directory (quota)	/home/z1234567 (15 GiB)	/home/561/zz1234 (10 GiB)
Shared scratch directory	/srv/scratch/z1234567	/scratch/aa00/zz1234
On-node fast scratch directory	\$TMPDIR (in /scratch)	\$PBS_JOBFS (in /jobfs)
Global data directory	/srv/scratch/group123	/g/data/aa00
I/O intensive file system	None	/iointensive
Home directory backups	Daily	Hourly, in \$HOME/.snapshot
Tape backup system	None	Via the mdss command

Module system

- Both Katana and Gadi use **module load / unload**
- Names of applications may be subtly different!
 - e.g., `r/4.3.1` → `R/4.3.1`; `python/3.11.3` → `python3/3.11.7`
- Different versions may be installed
- Some applications may require **#PBS -l software=** directives and additional modules
 - e.g., `#PBS -l software=matlab_unsw`
`module load matlab_licence/unsw`
- Use **module avail** on both systems to check
 - e.g., `module avail python`

#PBS directives

	Katana	Gadi
Charge resources from this project		#PBS -P <i>prj</i>
Which queue to submit to		#PBS -q <i>queue</i>
Request <i>m</i> CPUs over <i>n</i> nodes	#PBS -l select= <i>n</i> :ncpus= <i>m</i> :...	
Request <i>n</i> CPU cores in total		#PBS -l ncpus= <i>n</i>
Request <i>n</i> GPUs in total		#PBS -l ngpus= <i>n</i>
How much memory is required	#PBS -l select=...:mem= <i>size</i> GB	#PBS -l mem= <i>size</i> GB
How much time is required	#PBS -l walltime= <i>hh:mm:ss</i>	#PBS -l walltime= <i>hh:mm:ss</i>
Use software licence <i>licname</i>		#PBS -l software= <i>licname</i>
Use specific file systems		#PBS -l storage=scratch/ <i>prj</i> +gdata/ <i>prj</i>
Use on-node fast scratch (<i>size</i> GiB)		#PBS -l jobfs= <i>size</i> GB
Use I/O intensive storage (<i>size</i> TiB)		#PBS -l iointensive= <i>size</i>

Common job queues

Queue	CPUs + GPUs	Memory	Charge	Comments
normal	2×24	190 GiB	2.0	Intel Xeon Cascade Lake nodes
express	2×24	190 GiB	6.0	Intel Xeon Cascade Lake; high-priority queue
normalsr	2×52	500 GiB	2.0	Intel Xeon Sapphire Rapids nodes
normalsl	2×16	192 GiB	1.5	Older Intel Xeon Skylake nodes
normalbw	2×14	128/256 GiB	1.25	Old Intel Xeon Broadwell nodes
gpuvolta	2×24 + 4×V100	382 GiB	3.0	Nvidia Tesla V100; must use multiples of 12 cores
dxga100	2×64 + 8×A100	2000 GiB	4.5	Nvidia DGX A100; must use multiples of 16 cores
hugemem	2×24	1470 GiB	3.0	Intel Xeon Cascade Lake nodes
megamem	2×24	2990 GiB	5.0	Intel Xeon Cascade Lake nodes
hugemembw	2×14	1020 GiB	1.25	Broadwell nodes; must use multiples of 7 cores
megamembw	4×8	3000 GiB	1.25	Broadwell nodes; must use 32 or 64 cores
copyq	2×24	192 GiB	2.0	For file copying; must use one core only

Example job script

On Katana

```
#!/bin/bash

#PBS -M z1234567@unsw.edu.au
#PBS -m abe
#PBS -l walltime=01:00:00
#PBS -l select=1:ncpus=4:mem=8GB

cd $PBS_0_WORKDIR
module load python/3.11.3
./myjob.py /srv/scratch/z1234567/data.csv
```

Submit with **qsub ./myjob.sh**

On Gadi

```
#!/bin/bash

#PBS -M z1234567@unsw.edu.au
#PBS -m abe
#PBS -l walltime=01:00:00
#PBS -l ncpus=4
#PBS -l mem=8GB
#PBS -P aa00
#PBS -l storage=scratch/aa00

cd $PBS_0_WORKDIR
module load python3/3.11.7
./myjob.py /scratch/aa00/zz1234/data.csv
```

Submit with **qsub -q normalsl ./myjob.sh**

Resources

	Katana	Gadi
Responsible organisation	UNSW ResTech	NCI
Problems and requests	restech.support@unsw.edu.au	help@nci.org.au
Response time	~ 2–4 business days	~ 1–2 business days
System documentation	https://docs.restech.unsw.edu.au/	https://opus.nci.org.au/
Training courses	Via UNSW ResTech Community	Via NCI User Training
Check resource usage	Not needed	nci_account command
Check disk quota usage	disk-usage command	quota -s and lquota commands

Keep in contact!

Website: <https://unsw.sharepoint.com/sites/Restech>

Email: restech@unsw.edu.au

Location: [Level 3, June Griffith Building \(F10\)](#)

Teams: https://unsw.to/restech_teams

